nature biotechnology

Lipid production in *Nannochloropsis gaditana* is doubled by decreasing expression of a single transcriptional regulator

Imad Ajjawi[1], John Verruto[1], Moena Aqui[1], Leah B Soriaga[1], Jennifer Coppersmith[1], Kathleen Kwok[1], Luke Peach[1], Elizabeth Orchard[1], Ryan Kalb[1], Weidong Xu[1], Tom J Carlson[1,2], Kristie Francis[1], Katie Konigsfeld[1], Judit Bartalis[1], Andrew Schultz[1], William Lambert[1], Ariel S Schwartz[1], Robert Brown[1] & Eric R Moellering[1]

Lipid production in the industrial microalga *Nannochloropsis gaditana* exceeds that of model algal species and can be maximized by nutrient starvation in batch culture. However, starvation halts growth, thereby decreasing productivity. Efforts to engineer *N. gaditana* strains that can accumulate biomass and overproduce lipids have previously met with little success. We identified 20 transcription factors as putative negative regulators of lipid production by using RNA-seq analysis of *N. gaditana* during nitrogen deprivation. Application of a CRISPR–Cas9 reverse-genetics pipeline enabled insertional mutagenesis of 18 of these 20 transcription factors. Knocking out a homolog of fungal $Zn(II)_2Cys_6$-encoding genes improved partitioning of total carbon to lipids from 20% (wild type) to 40–55% (mutant) in nutrient-replete conditions. Knockout mutants grew poorly, but attenuation of $Zn(II)_2Cys_6$ expression yielded strains producing twice as much lipid (~5.0 g m^{-2} d^{-1}) as that in the wild type (~2.5 g m^{-2} d^{-1}) under semicontinuous growth conditions and had little effect on growth.

Since the early 1970s, phototrophic microalgal-derived biodiesel has been hailed as a potential supplement to petroleum-based transport fuels[1–4]. For economic feasibility, several technological hurdles must be overcome, two of which are biological. First, the photosynthetic efficiency of algal cultures in converting light energy and CO_2 to biomass requires improvement. Second, carbon (C) partitioning to lipids must be maximized without affecting biomass productivity. Although improvements in lipid productivity have been reported for the green alga *Chlamydomonas reinhardtii*[5–7] and the diatom species *Thalassiosira pseudonana*[8] and *Phaeodactylum tricornutum*[9,10], these three species are model organisms with native lipid productivities far lower than those of *Nannochloropsis* species[11–13]. In fact, *N. gaditana* has a lipid-production capacity up to sixfold higher than those of these model organisms[11]. *Nannochloropsis* strains have been cultivated in outdoor photobioreactors and raceway ponds[14–16], and have been shown to outperform other industrial production candidates such as *Chlorella* sp. or *Tetraselmis* sp. in terms of lipid productivity[13,15]. *Nannochloropsis* sp. can produce a lipid content up to ~60% of their ash-free dry weight[15], thus exceeding the native production of the heterotrophic industrial yeast *Yarrowia lipolytica*, which accumulates lipids from glucose to up to ~36% of its dry weight. *Y. lipolytica* can produce 60% lipid through feeding with exogenous fatty acids or through metabolic engineering[17,18].

Despite decades of research, previous genetic-engineering approaches have failed to yield substantial improvements in lipid productivity for any *Nannochloropsis* species. A lack of efficient genome-editing tools for industrially relevant strains[14] and an inability to increase lipid productivity without decreasing growth rate have confounded efforts to engineer this genus[19].

Here, we report a high-efficiency CRISPR–Cas9 reverse-genetics pipeline for *N. gaditana* (CCMP1894), which we applied to identify a transcriptional regulator of lipid accumulation, $Zn(II)_2Cys_6$ (ZnCys). Modulation of ZnCys expression enabled us to double the strain's lipid productivity while retaining its ability to grow and fix CO_2 at levels nearly equivalent to those of the wild type (WT) strain under dense semicontinuous culture.

RESULTS

Transcriptional profiling during nitrogen starvation

Industrial strains of the genus *Nannochloropsis* accumulate high levels of the storage lipid triacylglycerol (TAG) when they are starved of nitrogen (N)[20,21]. Under these conditions, broad transcriptional responses occur that are correlated with TAG accumulation, and underlying changes in gene expression in different metabolic pathways resulting in TAG synthesis have been proposed[11,21,22]. To identify transcriptional regulators of lipid accumulation that might be engineered to increase lipid production, we used RNA-seq to analyze the transcriptional response of *N. gaditana* cells subjected to N deprivation (−N) by centrifugation and resuspension in medium lacking N, and compared the response with that of cells resuspended in N-replete medium (+N, 8.8 mM NO_3^-) (Online Methods). Lipid accumulation (measured as the amount of fatty acid methyl esters (FAME)) was indistinguishable between cells grown in −N and +N conditions at 3 h after adjustment of N content in the growth medium,

[1]Synthetic Genomics Inc., La Jolla, California, USA. [2]Deceased. Correspondence should be addressed to I.A. (iajjawi@syntheticgenomics.com) or E.R.M. (emoeller@syntheticgenomics.com).

Received 8 November 2016; accepted 5 April 2017; published online 19 June 2017; doi:10.1038/nbt.3865

but at 10 h, FAME accumulation had doubled in the –N cultures (**Fig. 1a**). We selected the 3-h time point for mRNA sequencing because we assumed that transcriptional changes would precede metabolic changes.

Transcriptome analysis with RNA-seq of –N and +N cultures at 3 h revealed 1,064 differentially expressed genes (>0.5 $\log_2$ fold change (–N/+N), false discovery rate (FDR) <0.01; **Fig. 1b**). Of these, 363 were upregulated, and 701 were downregulated during N depletion. Using the Plant Transcription Factor Database as a reference[23], we identified 20 putative transcription factors (TFs) (**Supplementary Table 1**) that were downregulated in –N conditions. No TFs were upregulated at 3 h in –N conditions.

Candidate regulators of lipid accumulation

We hypothesized that a negative regulator of lipid accumulation would be downregulated in –N conditions. To analyze our set of 20 candidate TFs by reverse genetics, we needed to develop efficient genetic engineering tools for *N. gaditana*. We developed a CRISPR–Cas9 reverse-genetics pipeline, because a large amount of work would have been required to build 20 RNA interference (RNAi) constructs, and RNAi is not always able to produce knockdown lines with stable phenotypes[24]. First, we developed a high-efficiency Cas9-expressing *N. gaditana* 'editor line' (Ng-Cas9+; Online Methods). Eighteen of the 20 downregulated TFs were then disrupted in Ng-Cas9+ by insertion of a hygromycin resistance (HygR) cassette in the first half of the coding sequence of each TF-encoding gene, to produce loss-of-function insertional knockout (KO) mutants. For two of the targets, we were unable to identify insertional KOs (**Supplementary Table 1**). Correct insertion of cassettes was confirmed through PCR genotyping (Online Methods and **Supplementary Table 1**). Although Cas9-mediated mutagenesis of a single *Nannochloropsis* nitrate reductase–encoding gene has been reported[25], our method is superior in several ways. Wang *et al.*[25] needed to perform PCR analysis of 300 colonies to identify two nitrate reductase–KO strains, whereas our method generated two or more Cas9-mediated insertional-disruption strains in 32 or fewer transformants for each of 18 gene targets (**Supplementary Table 1**). Our method is more rapid, and we were able to identify gene KOs ~21 d after coelectroporation of guide RNAs (gRNAs) and selectable-marker-insert DNA, in contrast to the previously reported ~80 d required to isolate mutants in the workflow of Wang *et al.*[25]. Finally, we used *in vitro*–synthesized gRNAs and a stable Cas9-expressing editor line (NgCas9+), thereby avoiding the need to produce a separate gRNA-expressing plasmid for each gene target.

ZnCys transcription-factor mutants accumulate lipid

Two independent mutants per TF-encoding locus were screened for increased partitioning to lipids by assessing the ratio of FAME to total organic C (TOC) at multiple time points during a batch growth curve (Online Methods and **Supplementary Fig. 1a**). Cas9-mediated insertional KOs with a HygR-cassette insertion (**Fig. 1c**,**d**) in the coding sequence of locus Naga_100104g18 (referred to hereafter as *ZnCys*) encoding a predicted $Zn(II)_2Cys_6$ binuclear cluster domain and nuclear localization signal (Pfam PF00172; **Fig. 2a**) had a threefold increase in lipid accumulation compared with that of the WT, as assessed on the basis of the FAME/TOC ratio (**Fig. 1e** and **Supplementary Fig. 1b–d**). The presence of lipid droplets (**Fig. 1f**) and similar fatty acid profiles and TAG-accumulation phenotypes (**Supplementary Fig. 1e**,**f**) revealed that the lipid-accumulation phenotype in *ZnCys*-KO mutants mimicked that of N-starved WT *N. gaditana*. C partitioning to lipid in the *ZnCys*-KO lines was increased ~200% compared with that of the WT, and ~55% of TOC was allocated to fatty acids, thus corresponding to a FAME/TOC ratio of ~0.75.

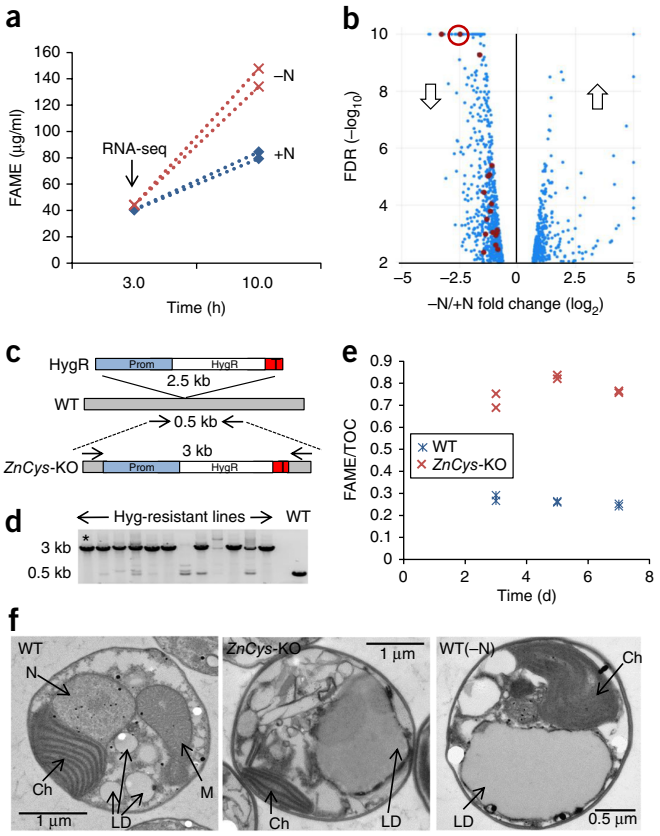

Figure 1 Engineering a *Nannochloropsis* strain with elevated lipid content. (**a**) FAME accumulation of WT *Nannochloropsis* cultures grown in N-replete (+N) and N-depleted (–N) conditions. FAME were quantified by gas chromatography with a flame ionization detector at 3 h and 10 h. Cells were grown under light limitation (optical density 1.25) and were resuspended in +N or –N medium at the beginning of the photoperiod. The samples at the 3-h time point were subjected to transcriptomic analysis (RNA-seq). Data for cell culture duplicates are shown. (**b**) Fold changes ($\log_2$) of genes (circles) differentially expressed (FDR <0.01) in –N cells relative to +N cells. Putative TFs disrupted by Cas9-mediated insertions are highlighted in red. Genes at the positive side of the *x* axis are upregulated (arrow pointing up), and genes at the negative side are downregulated (arrow pointing down). The *ZnCys* data point is encircled in red. (**c**) Schematic representation of a Cas9-mediated insertion of the HygR cassette into the *ZnCys* locus. The HygR cassette consisted of a promoter (prom) driving the HygR gene followed by a terminator (T). The resulting mutant genotype (*ZnCys*-KO) was identified by PCR with primers flanking the insertion (arrows). The cartoon is not to scale. (**d**) PCR genotyping of several hygromycin-resistant colonies transformed with a gRNA designed to target the *ZnCys* locus. The presence of a 3-kb band indicates insertion at the intended locus, whereas a 0.5-kb band indicates an intact WT locus. *ZnCys*-KO, indicated with an asterisk, was further characterized in **e** and **f**. (**e**) FAME/TOC ratios of *ZnCys*-KO and WT screened in batch mode (Online Methods) on nitrate-based standard medium (SM-NO_3^-). Data for cell culture duplicates are shown. (**f**) Representative electron micrographs of WT and *ZnCys*-KO cultured in SM-NO_3^- and WT in –N. Approximately 100 cells were imaged for each strain, from two replicates sampled from two independent experiments. N, nucleus; Ch, chloroplast; LD, lipid droplet; M, mitochondria.

Fine-tuning *ZnCys* expression to optimize lipid production

Although C partitioning to lipid increased in *ZnCys*-KO lines, we found an ~85% decrease in TOC productivity during batch growth (**Supplementary Fig. 1d** and **Fig. 2c**), which translated to a lipid

productivity lower than that of the WT under semicontinuous production. Hence, we investigated whether engineering decreased expression of *ZnCys* could be used to optimize lipid and TOC productivity. We hypothesized that insertions in the *ZnCys* 5′ and 3′ untranslated regions (UTRs) might result in attenuation of the encoded transcript by altering the transcription rate, the stability of the mRNA and/or the translational efficiency. To achieve these insertions, we performed Cas9-mediated insertion of the HygR cassette into *ZnCys* 5′- and 3′-UTR regions. We also used RNAi[26] (**Fig. 2a**) to validate whether downregulation of *ZnCys* transcript levels could optimize lipid productivity.

Using these approaches, we identified three *ZnCys*-attenuated lines that were suitable for further characterization: *ZnCys*-RNAi-7, which had an ~70% decrease in *ZnCys* mRNA, and two Cas9-mediated *ZnCys* UTR-insertion lines, *ZnCys*-BASH-3 in the 5′ UTR and *ZnCys*-BASH-12 in the 3′ UTR, which had ~20% and ~50% decreases in *ZnCys* mRNA, respectively (**Fig. 2a**,**b**). In batch growth on NO_3^-, all three *ZnCys*-attenuated strains had increased C partitioning to lipid (as FAME/TOC), but this increase was lower than that observed in the *ZnCys*-KO (**Fig. 2c**). TOC productivities in these *ZnCys*-attenuated lines were almost equivalent to those of the WT (with an ~18% decrease in the *ZnCys*-RNAi-7 line being the most severe) and were substantially improved compared with the ~85% deficit observed in the *ZnCys*-KO mutant (**Fig. 2c** and **Supplementary Fig. 2**).

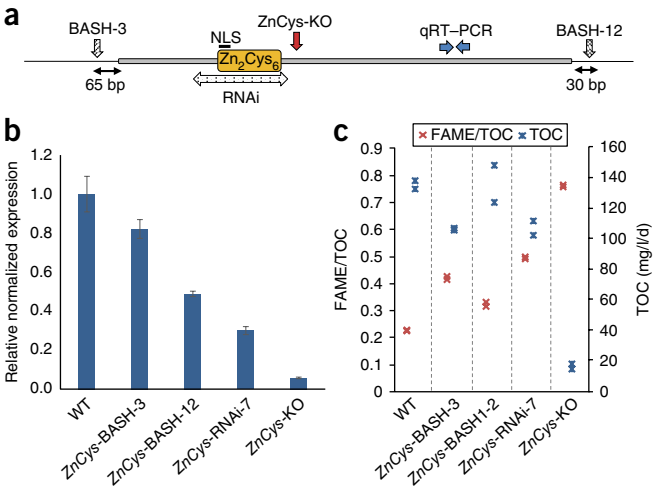

Figure 2 Generation of attenuated *ZnCys* alleles. (**a**) Modular structure and salient features of the *ZnCys* locus. NLS, nuclear localization sequence; Zn_2Cys_6, $Zn(II)_2Cys_6$ binuclear cluster domain (Pfam PF00172). The approximate location of the insertion in the original *ZnCys*-KO mutant is indicated with a red arrow. Patterned arrows indicate locations of successful Cas9 insertional mutants in the putative 5′-UTR (BASH-3, ~65 bp from the predicted start site) and 3′-UTR (BASH-5, approximately 30 bp from the predicted stop codon) regions. Blue arrows show the approximate location of the qRT–PCR primers used for assessing gene expression levels in **b**. The RNAi hairpin (dotted double arrow) designed to silence *ZnCys* spans the conserved Zn_2Cys_6 domain. The figure is not to scale. (**b**) Steady-state mRNA levels of the *ZnCys* locus in *ZnCys*-attenuated lines (*ZnCys*-BASH-3, *ZnCys*-BASH-12, *ZnCys*-RNAi-7 and *ZnCys*-KO) relative to WT grown on SM-NO_3^-, as determined by qRT–PCR. Expression levels were normalized to the control gene Naga_100004g25 and were calculated relative to WT with the ΔΔCt method. Error bars, s.e.m. for 3 technical replicates. (**c**) TOC productivity (blue) and FAME/TOC (red) values of *ZnCys*-mutant lines, assessed in batch mode in nitrate-replete medium (SM-NO_3^-). Individual data points used to calculate FAME/TOC and biomass productivity averages are shown in **Supplementary Figure 2**. Data from duplicate cell cultures are shown.

Sustained semicontinuous lipid production is doubled in *ZnCys*-RNAi-7

To accurately assess the extent and sustainability of lipid-productivity increases observed in the *ZnCys*-attenuated lines, we developed a semicontinuous productivity assay by using a 30% daily dilution of cultures, which yielded the optimum WT TOC productivity with an irradiance profile mimicking that of an average spring day in southern California (Online Methods). *ZnCys*-BASH-3, *ZnCys*-BASH-12 and *ZnCys*-RNAi-7 all had substantial increases in aerial FAME productivity (up to ~103% in *ZnCys*-RNAi-7) with only minimal TOC productivity decreases of ~5–15% compared with those of both WT *N. gaditana* and the Ng-Cas9+ parental strain (**Fig. 3a** and **Supplementary Figs. 3** and **4**). Under these conditions, the *ZnCys*-KO strain was washed out of the culture, owing to its inability to maintain growth under a 30% dilution scheme, as seen in the continued daily decline in TOC levels (**Supplementary Fig. 3**).

Flow cytometry cell counts revealed a decrease in cell number for the *ZnCys*-attenuated lines, including an ~38% decrease in *ZnCys*-RNAi-7 (**Supplementary Fig. 4c**). We assessed the *ZnCys*-BASH-12 and *ZnCys*-RNAi-7 lines for changes in C allocation to carbohydrates and proteins, which, together with FAME, accounted for ~75% of the TOC. The results revealed a specific decrease in protein content in the *ZnCys*-modified lines (~45–50%) compared with that in WT *N. gaditana*, on the basis of TOC (**Fig. 3b**). On a 'per cell' basis, the decrease in protein content was more modest, at ~15% for *ZnCys*-RNAi-7.

Moreover, *ZnCys*-KO had the highest C/N ratios, starting at ~10 and reaching a high of ~25 during the end of the run as TOC levels substantially decreased. *ZnCys*-RNAi had intermediate levels, averaging ~13 for the duration of the entire run, whereas the WT C/N ratio was ~7 (**Supplementary Fig. 3**). ^{14}C-based tracer analysis of photosynthetic capacity (P_{max}) at saturating irradiance revealed a decrease in the maximum rate of photosynthesis in *ZnCys*-BASH-12 compared with WT *N. gaditana* that was more severe on a TOC basis than on a per-cell basis (i.e., 36% versus ~14%; **Supplementary Table 2**).

Together, our findings indicated that under conditions of optimal semicontinuous aerial productivity, *ZnCys*-attenuated lines assimilated near-WT levels of C in the culture system with substantially fewer cells and less total protein, such that the decrease in protein per cell was moderate. Moreover, in dense semicontinuous algal cultures (i.e., self-shaded cultures attenuating ≥95% of incident irradiance), TOC productivities near those of WT were achieved in *ZnCys*-attenuated lines with substantially decreased photosynthetic capacity. A detailed characterization of the photophysiological phenotypes in *ZnCys*-attenuated lines is thus an important avenue for future research.

Transcriptional changes in *ZnCys*-RNAi-7

Experiments in which the *ZnCys*-KO strain was grown in medium supplemented with NH_4^+ rather than NO_3^- resulted in FAME and TOC productivities equivalent to those of WT *N. gaditana* (**Supplementary Fig. 5a–c**). The mutant phenotype of the *ZnCys*-KO strain was chemically complemented by NH_4^+ supplementation, thus suggesting that NO_3^- assimilation in *ZnCys* mutants is impaired.

We hypothesized that *ZnCys* mutants were unable to reduce NO_3^- to NH_4^+ efficiently (or at all) and/or were unable to import NO_3^-. To gain further insights into the mechanisms enabling *ZnCys*-attenuated strains to accumulate TAGs and sustain growth, we analyzed transcriptome profiles of *ZnCys*-RNAi-7 under semicontinuous growth

conditions by using RNA-seq (**Fig. 3c** and **Supplementary Fig. 4**). Using a fold-change cutoff of two and an FDR <0.05, we found 1,118 genes to be differentially expressed in *ZnCys*-RNAi-7 compared with WT *N. gaditana*. 790 genes were upregulated, and 328 were downregulated in the mutant strain. From Gene Ontology (GO) enrichment analysis (Online Methods), genes in the biological process 'photosynthesis, light harvesting' were overrepresented in the downregulated gene set. (**Supplementary Table 3** and Online Methods). The upregulated gene set was enriched in genes encoding components of protein synthesis (**Supplementary Table 4**). This result might indicate that the ~45% decrease in C partitioning to protein observed in *ZnCys*-RNAi-7 resulted in a compensatory response wherein transcripts encoding protein-biosynthesis machinery are upregulated (**Fig. 3b**). Other notable features of the transcriptome analysis were the downregulation of genes involved in N assimilation—including a nitrate transporter (NRT2), nitrite reductase (NiR), two glutamine synthetases (GS1 and GS2), an ammonium transporter (AMT1) and an enzyme involved in molybdenum-cofactor biosynthesis (MoeA or CNX1), an essential component for NR activity—in *ZnCys*-RNAi-7 (**Fig. 3c** and **Supplementary Table 5**). Although downregulation of N-assimilation genes is consistent with a decrease in protein content, it is distinct from the N-deprivation response in WT *N. gaditana*, in which N-assimilation genes tend to be upregulated[11]. Similarly, the upregulation of genes involved in protein synthesis in the *ZnCys*-attenuated strain is also distinct from N starvation, in which the same genes are downregulated[11,22]. Overall, the transcriptional profiles of *ZnCys*-RNAi during N depletion appeared to be substantially different from published profiles of N-starved *N. gaditana*, although downregulation of photosynthesis-related genes was a shared response[11,21,22].

GO-category analysis did not reveal an enrichment in lipid-biosynthesis genes, but we identified 22 genes related to glycerolipid biosynthesis that were upregulated in *ZnCys*-RNAi-7 (**Supplementary Table 6**). These genes included those encoding six fatty acid desaturases, elongases, lipases and acyltransferases of unknown substrate specificity, and the lipid droplet surface protein (LDSP)[27]. *Nannochloropsis* TAGs are primarily composed of C16:0 and C16:1 fatty acids[28]. The content of polyunsaturated fatty acids, such as C20:5, decreased in *ZnCys*-attenuated lines from 22.7% ± 0.2% to 6.2% ± 0.3% (mean ± s.d.), thus suggesting that upregulation of desaturases and elongases may be a response to the shortage of these membrane building blocks. The upregulation of LDSP in *ZnCys*-RNAi-7 is consistent with its accumulation during N deprivation and its proposed role in TAG lipid-droplet biogenesis[27]. *ZnCys*-RNAi-7 did not show differential regulation of genes encoding diacylglycerol acyltransferases or other enzymatic steps leading to TAG biosynthesis, thus indicating that upregulation of genes in the TAG-biosynthetic pathway is not required for the lipid phenotype in *ZnCys*-attenuated lines.

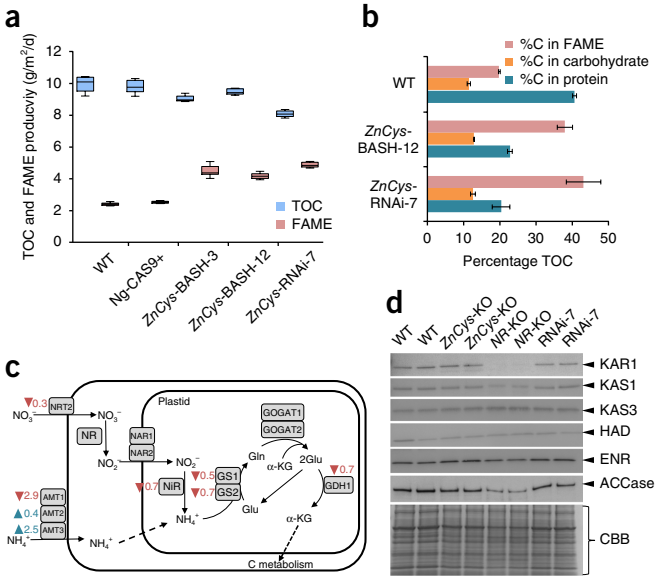

Figure 3 Productivity assessment and further characterization of *ZnCys*-mutant strains. (**a**) FAME (light red) and TOC (light blue) productivities (g/m^2/d) of *ZnCys* mutants (ZnCys-RNAi-7, ZnCys-BASH-12 and ZnCys-BASH-3) compared with their parental lines Ng-CAS9+ and WT. Cultures were grown in semicontinuous mode at a 30% daily dilution rate on SM-NO_3^- (details in Online Methods). Productivity values are presented as a box-and-whisker plot based on daily productivity measurements for 8 d from triplicate cell cultures ($n = 8$; **Supplementary Fig. 3**); whiskers, maximum and minimum data points; upper and lower bounds of the box, first and third quartiles; bar, median. (**b**) Biomass composition of WT, *ZnCys*-BASH-12 and *ZnCys*-RNAi-7 as percentage C measured in FAME, carbohydrate and protein. (**c**) Primary N-assimilation-pathway gene schematic depicting fold changes observed in *ZnCys*-RNAi-7 versus WT transcript abundance (from **Supplementary Table 5**). Enzymes and transporter-encoding genes are shown in boxes above reaction arrows. Arrowheads pointing up or down indicate transcripts up- or downregulated in *ZnCys*-RNAi-7, with FDR $<1.0 \times 10^{-2}$, and $\log_2$ fold changes are shown next to the arrowheads. NRT, nitrate transporter; AMT, ammonium transporter; NAR, nitrite transporter, NR, nitrate reductase; NiR, nitrite reductase; GS, glutamine synthase; GOGAT, glutamine oxoglutarate aminotransferase/glutamate synthase; GDH, glutamate dehydrogenase; Glu, glutamate; Gln, glutamine; α-KG, alpha-ketoglutarate. (**d**) Immunoblot analysis of enzymes in the FAS cycle (KAR1, ketoacyl-ACP synthases 1 and 3 (KAS1 and KAS3), hydroxyacyl-ACP dehydratase (HAD) and enoyl-ACP reductase (ENR)) and ACCase for duplicate cultures grown in batch mode on SM-NO_3^- (WT, *ZnCys*-KO and *ZnCys*-RNAi-7, shown as RNAi-7). *N. gaditana* identifiers and protein descriptions are available in **Supplementary Tables 6** and **7**. Coomassie brilliant blue stain (CBB) of a protein gel used for blotting is shown as a loading reference.

Inhibiting protein synthesis triggers lipid accumulation

We next assessed protein-level changes in the plastid lipid-biosynthetic machinery by using specific peptide antibodies against acetyl-CoA carboxylase (ACCase) and fatty acid synthesis (FAS)-pathway components, including ketoacyl-acyl carrier protein (ACP) reductase (KAR1), ketoacyl-ACP synthases 1 and 3, hydroxyacyl-ACP dehydratase and enoyl-ACP reductase (**Fig. 3d**). These enzymes were not expressed at higher levels in *ZnCys*-KO or *ZnCys*-RNAi-7 than in WT. In addition, a Cas9-mediated insertional-disruption mutant of nitrate reductase was engineered in the NgCas9+ editor line in a manner similar to that described for the 20 TFs targeted above. The resulting strain, named (*NR*-KO), was analyzed by first culturing in NH_4^+-containing medium, then subculturing in nitrate-based standard medium (SM-NO_3^-) medium to initiate batch growth (Online Methods). This procedure yielded a nitrogen-starved state by genetically inactivating nitrate assimilation. Western blot analysis of *NR*-KO cells under these growth conditions revealed decreases in ACCase and KAR1 proteins, as compared with the levels in WT, *ZnCys*-KO and *ZnCys*-RNAi-7 lines. These data suggested that under N-replete conditions, the plastid lipid-biosynthetic machinery is present at a capacity allowing for (at least) double the flux to fatty acid, given the ~100% increase in FAME productivity observed in *ZnCys*-RNAi-7. This result was consistent with the transcriptional changes reported in previous studies of N-deprived WT *N. gaditana* and *Nannochloropsis oceanica*. In those studies, pathways involved in providing C precursors to lipid biosynthesis

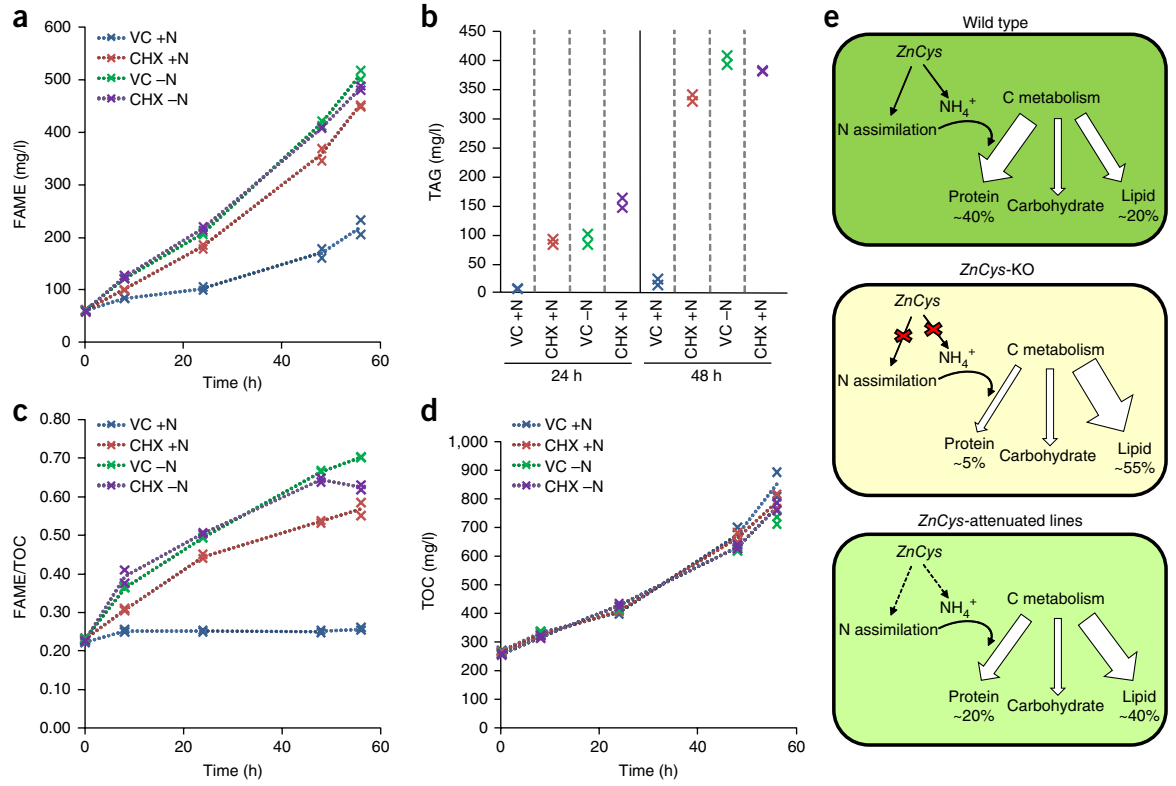

Figure 4 Effects of CHX treatment on FAME productivity and C partitioning to lipid under N-replete and N-deficient batch growth. (**a**–**d**) FAME/TOC ratios (**a**), TAG levels (**b**), FAME accumulation (**c**) and TOC accumulation (**d**) in WT cells treated with either 25 µg/ml CHX or vehicle control (VC, ethanol at 0.25% (vol/vol)) at 0 h. At 0 h, cells were harvested by centrifugation and resuspended in N-replete medium (+N; SM-NO_3^-) or N-free medium (–N). Shown are data from duplicate cell cultures; a dotted line connects the mean across time points. (**e**) Schematic of C partitioning to protein, lipid and carbohydrate in WT, *ZnCys*-KO and *ZnCys*-attenuated lines. Expression of *ZnCys* can be fine-tuned to increase C-allocation to lipid at the expense of protein, thus doubling lipid productivity.

were differentially expressed to a greater extent than 'core' lipid-biosynthetic pathways during nitrogen depletion[11,21,22].

We did not observe drastic changes in the relative ratios of the separated individual protein bands of WT and *ZnCys*-RNAi-7 in a Coomassie-stained gel loaded with equal amounts of extracted protein (**Fig. 3d**). Because there was an ~45% decrease in total protein per TOC in *ZnCys*-attenuated lines (**Fig. 3b**), it is possible that selective maintenance and/or removal of specific proteins may have occurred, as previously observed in the nitrogen-starvation response of mixotrophically grown *C. reinhardtii*[29].

We investigated whether *de novo* synthesis of proteins during –N batch growth was required for lipid induction in WT *N. gaditana* cultures. Treatment of –N cultures with the cytosolic ribosomal inhibitor cycloheximide (CHX) revealed that lipid accumulation occurred at the same level as that in untreated controls (**Fig. 4**). Furthermore, CHX treatment of N-replete cultures was sufficient to induce partitioning to lipids (FAME/TOC) and TAG accumulation (**Fig. 4a**,**b**) at a magnitude similar to that under N deprivation, and there was no significant decrease in TOC productivity (**Fig. 4d**) during a 56-h time course. These data suggested that *N. gaditana* diverts excess C to TAG when protein synthesis is restricted, in agreement with the observed reallocation of C during N deprivation.

DISCUSSION

We developed a CRISPR–Cas9 reverse-genetics pipeline and used it to identify a regulator of lipid accumulation named ZnCys in *N. gaditana* (CCMP1894). We modulated expression of this regulator through Cas9-mediated insertional attenuation in the 5′ UTR and RNAi, and found that for *ZnCys*-RNAi-7, lipid productivity was doubled in the absence of changes in the expression of FAS or components of the glycerolipid-assembly pathway leading to TAG synthesis. Downregulation of multiple genes involved in N assimilation appears to restrict flux to protein synthesis, thus resulting in the allocation of C into storage TAGs (**Fig. 4e**). A number of other transcriptional changes arising from the fine-tuning of *ZnCys* expression may also contribute to the performance of *ZnCys*-attenuated lines with respect to lipid- and biomass-productivity phenotypes. These results do not support the 'algal lipid trigger' hypothesis, in which transcriptional upregulation of fatty acid synthase and TAG-assembly proteins underpin lipid induction[21,22,30,31].

Our results provide some insights into the mechanisms underlying regulation of lipid accumulation in *N. gaditana*. First, our CHX experiments in –N cultures suggest that transcript-level changes in response to N deprivation that result in altered abundance in nuclear-encoded proteins may not be required for induced lipid biosynthesis. Instead, it seems more likely that the enzymes for TAG synthesis are present in sufficient quantities before N depletion, in agreement with the lack of transcript-level changes observed for FAS and TAG-assembly proteins, and the lack of changes in FAS-protein abundance (**Fig. 4d**) in *ZnCys*-attenuated lines.

We cannot exclude the possibility that transcriptional changes during –N conditions may be required for optimal, or continued, lipid accumulation. In addition, the induced lipid accumulation observed in CHX-treated N-replete cultures, together with the abil-

ity to modulate FAME/TOC levels in a CHX-dose-dependent manner (**Supplementary Fig. 6**), may indicate that restricting flux to protein synthesis is sufficient to induce lipid accumulation in *N. gaditana*.

Maximizing C partitioning to lipid while maintaining high biomass productivity is crucial if microalgal-derived biodiesel is to be commercially viable. As has been reported in industrial oleochemical heterotrophic hosts such as *Yarrowia*[17,32], we showed that lipid productivity can be substantially improved in *N. gaditana*. Using a microalga to produce lipids offers the potential advantages of being able to phototrophically convert CO_2 to lipids without relying on agriculturally derived sugars, thus mitigating the demand for arable land and freshwater[33,34].

Our findings represent a step toward understanding and controlling lipid production in algae. This ability to control algal lipid production might eventually enable the commercialization of microalgal-derived biofuels.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was funded by ExxonMobil and Synthetic Genomics, Inc. We thank A. Withrow (Center for Advanced Microscopy at Michigan State University) for producing the TEM images and C. Packard, B. Scherer, E. Wang and the rest of the analytical team at SGI for processing FAME and TOC samples. This work is dedicated to our colleague Tom Carlson, who passed away during the preparation of this manuscript.

AUTHOR CONTRIBUTIONS

I.A. and E.R.M. conceived the study and designed experiments. R.B. provided technical advice. E.O. and R.K. designed the productivity assays. L.B.S. and A.S.S. performed computational and bioinformatics analyses. M.A., J.V., J.C., L.P., J.B., A.S., W.X., T.J.C., K.F., W.L., K. Kwok and K. Konigsfeld performed the experiments. I.A. and E.R.M. wrote the manuscript with support from all authors.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Goncalves, E.C., Wilkie, A.C., Kirst, M. & Rathinasabapathi, B. Metabolic regulation of triacylglycerol accumulation in the green algae: identification of potential targets for engineering to improve oil yield. *Plant Biotechnol. J.* **14**, 1649–1660 (2016).
2. Hu, Q. *et al.* Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant J.* **54**, 621–639 (2008).
3. Vasudevan, V. *et al.* Environmental performance of algal biofuel technology options. *Environ. Sci. Technol.* **46**, 2451–2459 (2012).
4. Wijffels, R.H. & Barbosa, M.J. An outlook on microalgal biofuels. *Science* **329**, 796–799 (2010).
5. Ngan, C.Y. *et al.* Lineage-specific chromatin signatures reveal a regulator of lipid metabolism in microalgae. *Nature Plants* **1**, 15107 (2015).
6. Goold, H.D. *et al.* Whole genome re-sequencing identifies a quantitative trait locus repressing carbon reserve accumulation during optimal growth in *Chlamydomonas reinhardtii*. *Sci. Rep.* **6**, 25209 (2016).
7. Schulz-Raffelt, M. *et al.* Hyper-accumulation of starch and oil in a *Chlamydomonas* mutant affected in a plant-specific DYRK kinase. *Biotechnol. Biofuels* **9**, 55 (2016).
8. Trentacoste, E.M. *et al.* Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth. *Proc. Natl. Acad. Sci. USA* **110**, 19748–19753 (2013).
9. Daboussi, F. *et al.* Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nat. Commun.* **5**, 3831 (2014).
10. Levitan, O., Dinamarca, J., Zelzion, E., Gorbunov, M.Y. & Falkowski, P.G. An RNA interference knock-down of nitrate reductase enhances lipid biosynthesis in the diatom *Phaeodactylum tricornutum*. *Plant J.* **84**, 963–973 (2015).
11. Radakovits, R. *et al.* Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropis gaditana*. *Nat. Commun.* **3**, 686 (2012).
12. Griffiths, M.J. & Harrison, S.T.L. Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *J. Appl. Phycol.* **21**, 493–507 (2009).
13. Ma, X.N., Chen, T.P., Yang, B., Liu, J. & Chen, F. Lipid production from *Nannochloropsis*. *Mar. Drugs* **14**, 61 (2016).
14. Kilian, O., Benemann, C.S., Niyogi, K.K. & Vick, B. High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp. *Proc. Natl. Acad. Sci. USA* **108**, 21265–21269 (2011).
15. Rodolfi, L. *et al.* Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnol. Bioeng.* **102**, 100–112 (2009).
16. Sukenik, A. *et al.* Photosynthetic perfromance of outdoor *Nannochloropsis* mass cultures under a wide range of environmental conditions. *Aquat. Microb. Ecol.* **56**, 297–308 (2009).
17. Blazeck, J. *et al.* Harnessing *Yarrowia lipolytica* lipogenesis to create a platform for lipid and biofuel production. *Nat. Commun.* **5**, 3131 (2014).
18. Friedlander, J. *et al.* Engineering of a high lipid producing *Yarrowia lipolytica* strain. *Biotechnol. Biofuels* **9**, 77 (2016).
19. Zienkiewicz, K., Du, Z.Y., Ma, W., Vollheyde, K. & Benning, C. Stress-induced neutral lipid biosynthesis in microalgae: molecular, cellular and physiological insights. *Biochim. Biophys. Acta* **1861** 9 Pt B, 1269–1281 (2016).
20. Boussiba, S.V.A., Cohen, Z., Avissar, Y. & Richmond, A. Lipid and biomass production by the halotolerant microalga *Nannochloropsis salina*. *Biomass* **12**, 37–47 (1987).
21. Li, J. *et al.* Choreography of transcriptomes and lipidomes of nannochloropsis reveals the mechanisms of oil synthesis in microalgae. *Plant Cell* **26**, 1645–1665 (2014).
22. Corteggiani Carpinelli, E. *et al.* Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Mol. Plant* **7**, 323–335 (2014).
23. Pérez-Rodríguez, P. *et al.* PInTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* **38**, D822–D827 (2010).
24. Cerutti, H., Ma, X., Msanne, J. & Repas, T. RNA-mediated silencing in algae: biological roles and tools for analysis of gene function. *Eukaryot. Cell* **10**, 1164–1172 (2011).
25. Wang, Q. *et al.* Genome editing of model oleaginous microalgae *Nannochloropsis* spp. by CRISPR/Cas9. *Plant J.* **88**, 1071–1081 (2016).
26. Kennerdell, J.R. & Carthew, R.W. Heritable gene silencing in *Drosophila* using double-stranded RNA. *Nat. Biotechnol.* **18**, 896–898 (2000).
27. Vieler, A., Brubaker, S.B., Vick, B. & Benning, C. A lipid droplet protein of *Nannochloropsis* with functions partially analogous to plant oleosins. *Plant Physiol.* **158**, 1562–1569 (2012).
28. Xiao, Y., Zhang, J., Cui, J., Feng, Y. & Cui, Q. Metabolic profiles of *Nannochloropsis oceanica* IMET1 under nitrogen-deficiency stress. *Bioresour. Technol.* **130**, 731–738 (2013).
29. Schmollinger, S. *et al.* Nitrogen-sparing mechanisms in *Chlamydomonas* affect the transcriptome, the proteome, and photosynthetic metabolism. *Plant Cell* **26**, 1410–1435 (2014).
30. Gargouri, M. *et al.* Identification of regulatory network hubs that control lipid metabolism in *Chlamydomonas reinhardtii*. *J. Exp. Bot.* **66**, 4551–4566 (2015).
31. Hu, J. *et al.* Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae *Nannochloropsis*. *Sci. Rep.* **4**, 5454 (2014).
32. Xu, P., Qiao, K., Ahn, W.S. & Stephanopoulos, G. Engineering *Yarrowia lipolytica* as a platform for synthesis of drop-in transportation fuels and oleochemicals. *Proc. Natl. Acad. Sci. USA* **113**, 10848–10853 (2016).
33. Biddy, M.J. *et al.* The techno-economic basis for coproduct manufacturing to enable hydrocarbon fuel production from lignocellulosic biomass. *ACS Sustain. Chem.& Eng.* **4**, 3196–3211 (2016).
34. Davis, R., Aden, A. & Pienkos, P.T. Techno-economic analysis of autotrophic microalgae for fuel production. *Appl. Energy* **88**, 3524–3531 (2011).

ONLINE METHODS

Strain availability. *N. gaditana* strain CCMP1894 was obtained from CCMP, now known as the Bigelow National Center for Marine Algae and Microbiota (NCMA). All strains described in this manuscript were derived from this parental background. The Ng-Cas9+ strain can be requested from Synthetic Genomics, Inc., by contacting the corresponding authors, or through the ATCC. The cell line was authenticated by immunoblotting for Cas9 and confirming GFP expression by flow cytometry; it was not tested for mycoplasma contamination.

Media formulations. Standard medium (SM-NO_3^-) consisted of 35 g/l Instant Ocean, 10× F/2 trace metals and vitamins, and 0.361 mM NaH_2PO_4. The N source was 8.8 mM $NaNO_3$ for the semicontinuous assay and 15 mM $NaNO_3$ for the batch growth screen, both of which are described below. When cultures were supplemented with ammonium, 10 mM NH_4Cl was used as the N source and was buffered with 15 mM HEPES, pH 8.0; this medium is referred to as SM-NH_4^+/NO_3^- in the main text.

Batch growth assessment. Cultures (175 ml) were grown in 75 cm^2 rectangular tissue culture flasks positioned with their narrowest 'width' dimension against an LED light source. Culture flasks were masked with opaque white plastic to provide a 21.1-cm^2 rectangular opening allowing irradiance to reach the culture. Incident irradiance was programmed to mimic the light intensity of a 14-h light, 10-h dark diel period: a linear ramp up of irradiance from 0 to 1,200 µE over 4 h (increasing in 15-min intervals) was followed by 6 h at constant 1,200 µE and then a linear ramp down in irradiance from 1,200 to 0 µE over a 4-h period (**Supplementary Fig. 7**). The temperature of the cultures was regulated by a water bath set at 25 °C. Deionized H_2O was added to the cultures daily to replace evaporative losses. Cultures were inoculated at an $OD_{730\ nm}$ of 0.5, and 5-ml samples were typically harvested every other day for analytical measurements. Sampling time was kept constant at 30 min before the end of the light phase of the diel cycle.

Semicontinuous productivity assay. Rectangular tissue culture flasks (225 cm^2) were inoculated to an $OD_{730\ nm}$ of 0.9 in a final volume of 550 ml. The inoculum (typically 200 ml) was previously scaled up in batch mode in the same conditions to allow the cells to acclimate to the assay environment. Culture flasks were masked with opaque white plastic to provide a 31.5-cm^2 rectangular window allowing irradiance to reach the culture. The flasks were aligned with their width (narrowest dimension) against an LED light bank that was programmed with a light–dark cycle and a light profile that increased until 'solar noon' and then decreased until the end of the light period. The light profile was designed to mimic a spring day in southern California, with 14 h light and 10 h dark, with the light peaking at approximately 2,000 µE (**Supplementary Fig. 7**). The flasks included stir bars and had stoppers with inserted tubing connected with syringe filters for delivering CO_2-enriched air (1% CO_2, flow rate, 300 ml per min). The flasks were set in a water bath programmed to maintain a constant temperature of 25 °C on stir plates set to 575 r.p.m. during the assay period. Cultures were diluted daily at mid-day, when the light intensity was at its peak, by removing 30% of the volume (165 ml) and replacing it with the same volume of the assay medium plus an additional 10 ml of deionized water to compensate for evaporation. A 30% dilution rate was empirically determined as the most productive dilution rate for *Nannochloropsis*, because 50%, 30% and 15% daily dilutions resulted in average TOC productivities of 6.5, 9 and 8 g/m^2/d, respectively. Semicontinuous assays were typically run for 7–14 d. Daily lipid (FAME) and biomass (TOC) productivities were calculated from cultures that had reached steady-state standing-crop TOC and FAME density. Volumetric FAME and TOC productivities (in mg l^{-1} d^{-1}) were calculated by multiplying the volumetric FAME and TOC amounts by the 30% dilution rate. Aerial productivities (in g m^{-2} d^{-1}) were calculated by dividing the total productivity of the culture by the size of the aperture through which irradiance was permitted:

$$\text{volumetric productivity} \times \frac{0.551}{0.00315\ \text{m}^2} \times \frac{\text{g}}{1{,}000\ \text{mg}} = \frac{\text{g}}{\text{m}^2 \times \text{d}}$$

RNA extraction. Cell pellets were resuspended in 1.8 ml extraction solution (5 ml grinding buffer, 5 ml phenol, 1 ml 1-bromo-3-chloropropane and 20 µl mercaptoethanol, with grinding buffer comprising 9 ml of 1 M Tris, pH 8, 5 ml of 10% SDS, 0.6 ml of 7.5 M LiCl, and 450 µl 0.5 M EDTA in a final volume of 50 ml) and vortexed vigorously for 5 min at 4 °C in the presence of 200-µm zirconium beads. After centrifugation, 1 ml of 25:24:1 phenol extraction solution (25 ml phenol, pH 8.1, 24 ml 1-bromo-3-chloropropane and 1 ml isoamyl alcohol) was added to the aqueous phase in a separate tube. Tubes were shaken vigorously and centrifuged for 2 min at 21,000*g*. The extraction was repeated with 1 ml 1-bromo-3-chloropropane, and the resulting aqueous layer was treated with 0.356 volumes of 7.5 M LiCl to precipitate the RNA overnight at −20 °C. After LiCl precipitation, RNA pellets were resuspended in 50 µl H_2O, and RNA quality was assessed by on-chip gel electrophoresis with an Agilent 2100 Bioanalyzer, according to the manufacturer's instructions.

RNA-seq. Next-generation sequencing libraries were prepared from total RNA with a TruSeq Stranded mRNA Sample Prep Kit (Illumina), according to the manufacturer's instructions. For the nitrogen-replete and nitrogen-depleted comparison, TruSeq libraries were sequenced on an Illumina HiSeq platform (Ambry Genetics) to generate 100-bp paired-end reads (described in ref. 35). For the comparison of the *ZnCys*-RNAi-7 to WT, libraries were sequenced in house on an Illumina NextSeq platform to generate 75-bp paired-end reads. Trimmomatic software (http://www.usadellab.org/cms/?page=trimmomatic) was used to filter and trim reads on the basis of base-quality scores. Mapping of the trimmed-read libraries and expression estimation relative to a reference gene set were performed with RSEM[36]. The reference gene set comprised predicted coding sequences based on the *N. gaditana* B-31 genome annotation (http://www.nannochloropsis.org/page/ftp/). Differential expression analysis was performed with the R package edgeR[37]. Gene ontology annotation for the *N. gaditana* coding sequences were obtained from the UniProt-GOA database (ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/proteomes/357410.N_gaditana.goa). Gene set enrichment testing was performed with goseq in R to identify enriched GO categories among differentially expressed genes (FDR-adjusted *P* value <0.05) for a given comparison and fold-change direction[38].

Quantitative real-time PCR. Total RNA was reverse transcribed with an iScript Reverse Transcription Supermix kit (Bio-Rad). The enzyme mix Ssofast EvaGreen Supermix (Bio-Rad) was used for amplification of genes of interest. Primer and cDNA concentrations were as recommended by the manufacturer, and reactions were performed on a C1000 Thermal Cycler coupled with a CFX Real-time System (Bio-Rad). The primer sequences for *ZnCys* (Naga_100104g18) were F, 5′-atacaggaagcgtggttacag-3′ and R, 5′-gaagtattaaggactggccg-3′. qRT–PCR primers were evaluated for efficiency, and the $2^{-\Delta\Delta Ct}$ method was used to estimate gene expression normalized to that of a control gene (Naga_100004g25, with primer sequences F, 5′-ctctcctattgctttcccctcg-3′ and R, 5′-ctaccaacacctctacacttcc-3′) that was empirically determined to have a low coefficient of variation across different conditions.

Generation of the Cas9 editor line. The high-efficiency *N. gaditana* Cas9 editor line (Ng-Cas9+) was generated by random integration of a ZraI restriction enzyme–linearized vector containing expression cassettes for the selectable marker blasticidin deaminase (BSD), Cas9 and TurboGFP (Evrogen) (**Supplementary Fig. 8**). The BSD and Cas9 genes were optimized for *N. gaditana* codon usage, whereas TurboGFP was directly amplified from pTurboGFP-C purchased from Evrogen. The Cas9 coding sequence contained an N-terminal nuclear localization signal (NLS) and a FLAG epitope followed by a linker region, as described in ref. 39. The expression of all three genes was driven by endogenous promoters and terminators; the source genes are described in **Supplementary Table 8**. The Cas9 expression construct was assembled with a Gibson Assembly HiFi 1 Step Kit (Synthetic Genomics) into a minimal pUC-vector backbone; the confirmed DNA sequence of this plasmid is shown in GenBank format in **Supplementary Note 1**.

Transformation of the ZraI-linearized Cas9 expression construct by electroporation was conducted according to ref. 11, except that 1×10^9 cells were transformed in a 0.2-cm cuvette with a field strength of 7,000 V/cm, delivered with a Gene Pulser II (Bio-Rad). After overnight recovery in liquid SM-NO_3^- medium, cells were plated on SM-NO_3^- agar containing 100 mg l^{-1} blasticidin (Invivogen). Putative transformant colonies were replated on the same selective-medium plates, and after sufficient growth, an inoculation loop full

of cells was resuspended in liquid medium and analyzed on an Accuri C6 cytometer (BD Biosciences) for GFP fluorescence (**Supplementary Fig. 9a**). Stable transgenic Cas9 expression was confirmed in the Ng-Cas9+ line by western blotting with an anti-FLAG antibody from Sigma-Aldrich (cat. no. F1804; validation on manufacturer's website; **Supplementary Fig. 9b**). The Ng-Cas9+ line was assessed for potential differences in FAME or TOC productivity compared with those of WT under semicontinuous dilution growth mode and was found to be equivalent to WT in these parameters (**Fig. 3a** and **Supplementary Fig. 4**).

Generation of targeted insertional mutants in the Ng-Cas9+ editor line. Synthetic gRNAs were designed according Cho *et al.*[40] and Mali *et al.*[41]. Target regions were selected by searching for the presence of NGG (the protospacer-adjacent motif) in the gene of interest. The spacer sequence comprised 17–20 nt directly upstream of the protospacer-adjacent motif, as shown in **Supplementary Table 9**. gRNAs were synthesized *in vitro* according to ref. 40 and coelectroporated with a PCR-amplified expression cassette (pHygR) containing a codon-optimized hygromycin-resistance gene (HygR) driven by the endogenous EIF3_P promoter and FRD_T bidirectional terminator in inverted orientation (descriptions of DNA-element sources in **Supplementary Fig. 9b** and **Supplementary Table 8**). The confirmed DNA sequence of pHygR is shown in GenBank format in **Supplementary Note 2**. Approximately 1 µg each of gRNA and pHygR was added to the cuvette, and electroporation was conducted as described above. Selection of HygR transformants was as described above except that 500 mg l^{-1} hygromycin was added to agar plates instead of blasticidin. Targeted insertion of the pHygR fragment via NHEJ-mediated repair of the double-stranded DNA break catalyzed by Cas9 in loci targeted by gRNAs was assessed by colony PCR with primers flanking the gRNA target site by ~200 bp on either side. In this manner, PCR amplification of WT loci would result in ~400-bp products, whereas pHygR-targeted insertional mutants would result in ~2,800-bp products—thus accounting for the insertion of the 2,400-bp pHygR fragment. All insertional-mutant lines presented in this study were confirmed to have a single pHygR insert at the intended locus and to lack the WT PCR product. PCR products were sequence confirmed for all mutants, thus allowing the determination of insert orientation and potential loss of chromosomal and/or insert DNA, or the gain of small insertions generated during the NHEJ-mediated dsDNA-break-repair process (example colony PCR results in **Fig. 1d**). In general, loss-of-function KO mutants were generated by selecting a gRNA-target locus in the first half of an exonic coding sequence, whereas Cas9-derived insertional-attenuation mutants were created by targeting insertions to the promoter or untranslated regions of a given gene, as described in the main text (gRNA sequences in **Supplementary Table 9**).

Generation of *ZnCys*-RNAi lines. The RNAi plasmid designed to silence the *ZnCys* locus was constructed by using a minimal pUC vector as the backbone. The confirmed DNA sequence of the final RNAi plasmid is provided in GenBank format in **Supplementary Note 3**. Briefly, the sense and loop structures of the *ZnCys* hairpin were amplified from *N. gaditana* cDNA with primers F, 5′-gacagagacacacagggatcgtttaaacgatcagccacgacggctctcg-3′ and R, 5′-cactatgcaccttccccttgactcaccggttcgtggcactgctgttg-3′. These primers contained ~20-bp overlaps with the minimal pUC vector backbone to allow for cloning with a Gibson AssemblyHiFi 1 Step Kit (SGI-DNA). The antisense fragment was amplified with 5′-cacaacagcagtgccacgaaccgccccgactgcttttactgc-3′ and 5′-ctatgcaccttccccttgactcgtttaaacgatcagccacgacggctc-3′ primers and cloned into the vector containing the sense and loop fragments by digesting the vector with AgeI and by using the Gibson Assembly method (Synthetic Genomics). The final vector was digested with ScaI, and 1 µg of the pertinent linear fragment was transformed into WT *N. gaditana* as described above. 32 hygromycin-resistant transformant colonies were assessed for GFP expression as described above for the Cas9 editor line NgCas9+. Eight GFP-expressing constructs were screened for increased FAME/TOC, and the line *ZnCys*-RNAi-7 was selected because it displayed maximal FAME/TOC with little apparent decrease in TOC productivity.

Analytical methods (FAME, TOC and LC/GC–MS). FAME analysis was performed on 2-ml samples that were dried with a GeneVac HT-4X. To each dried pellet, the following were added: 500 µl of 500 mM KOH in methanol, 200 µl of tetrahydrofuran containing 0.05% butylated hydroxyltoluene, 40 µl of a 2 mg/ml C11:0 free fatty acid/C13:0 triglyceride/C23:0 fatty acid methyl ester internal standard mix and 500 µl of glass beads (diameter 425–600 µm). The vials were capped with open-top PTFE-septa-lined caps and placed in an SPEX GenoGrinder at 1,650 r.p.m. for 7.5 min. The samples were then heated at 80 °C for 5 min and allowed to cool. For derivatization, 500 µl of 10% boron trifluoride in methanol was added to the samples, which were then heated at 80 °C for 30 min. The tubes were allowed to cool, and 2 ml of heptane and 500 µl of 5 M NaCl were then added. The samples were vortexed for 5 min at 2,000 r.p.m. and finally centrifuged for 3 min at 1,000 r.p.m. FAME were quantified with gas chromatography with a flame ionization detector (GC–FID) on an Agilent 7980A instrument equipped with a DB-FFAP column (Agilent P/N 127-3212) with the C23:0 externally spiked FAME as a standard, by using methods similar to the American Oil Chemists' Society Methods Ce 1b-89 and Ce 1-62 (http://www.aocs.org/Methods/).

Total organic carbon (TOC) was determined by dilution of 2 ml of cell culture to a total volume of 20 ml with DI water. For each measurement, three injections into a Shimadzu TOC-Vcsj Analyzer were performed for determination of total carbon (TC) and total inorganic carbon (TIC). The combustion furnace was set to 720 °C, and TOC was determined by subtraction of TIC from TC. The four-point calibration range was from 2 p.p.m. to 200 p.p.m., corresponding to 20–2,000 p.p.m. for nondiluted cultures with a correlation coefficient of r^2 >0.999.

For LC–MS analysis of lipids, 2-ml samples of each culture were spun down at maximum speed for 5 min, the supernatants were removed, and total lipids were extracted from the pellets via a modified Bligh–Dyer extraction[42]. The bottom, lipid-containing layer was concentrated to dryness with an EZ-2 Genevac. The residue was dissolved in 1 ml of 1:1 chloroform/methanol containing 10 µg/ml BHT and 10 µg/ml TAG (17:1/17:1/17:1) as an internal standard. LC–MS was performed on an Agilent 1200 LC interfaced with an Agilent 6538 Q-ToF with a dual ESI source. 10 µl of the extract was injected onto a Halo C8 column, 2.7 µm, 4.6 × 150 mm (MacMod Analytics) and eluted in a gradient of mobile phase A (375:375:250:1 methanol/isopropanol/water/acetic acid + 10 mM ammonium acetate) and mobile phase B (500:375:125 acetonitrile/isopropanol/tetrahydrofuran + 10 mM ammonium acetate). The flow rate was 0.6 ml/min, and the following gradient elution was used: 0 min, 15% B; 1 min, 15% B; 15 min, 56% B; 19 min, 85% B; and 26 min, 85% B. The source parameters were as follows: gas temperature, 300 °C; drying gas, 10 l/min, nebulizer, 35 psig; capillary voltage, 4,000 V. The ToF scan rate was 1 spectrum/s. Individual TAGs were identified by accurate mass and quantified by summing the ion intensities of the individual TAGs and using a relative response factor of 1 compared with the ion intensity of the 10 µg/ml TAG(17:1/17:1/17:1) internal standard.

Total carbohydrate analysis was conducted on an ~0.7 mg TOC equivalent of cell culture purified by three rounds of centrifugation and washing with PBS, and concentrated to 0.5 ml in PBS. Acid hydrolysis was used to convert carbohydrates into their constituent monomers by the addition of 0.5 ml deionized H_2O and 1 ml 6 N HCl and [*U*-^{13}C]glucose and [*U*-^{13}C]galactose as internal standards at a final concentration of 50 µg/ml each. Samples were heated at 105 °C for 1 h in glass vials with PTFE-lined caps. 100-µl aliquots of the room-temperature-cooled samples that had been centrifuged at 3,000*g* for 1 min were dried in an EZ-2 Genevac, derivatized with MSTFA/TMCS and analyzed by GC–MS according to ref. 43. Internal ^{13}C-labeled standards were used to quantify the concentration of the major carbohydrate monomers, glucose and galactose, and to estimate the concentrations of less abundant sugars (arabinose, rhamnose, xylose and mannose). These concentrations were summed to yield a total saccharide concentration (in micrograms per milliliter) which was converted to the carbon content of total carbohydrates by a multiplication factor of 0.45 (i.e., ~45% of the carbohydrate mass represented by carbon). This value was divided by the amount of TOC detected in an identical aliquot of concentrated cell culture to estimate the percentage of carbon allocated to carbohydrate, as shown in **Figure 3b**.

Analysis of total amino acids was conducted by derivatization of whole amino acid hydrolysate to propoxycarbonyl propyl esters through a modified method, according to the manufacturer's instructions for the EZ:faast kit from Phenomenex. Briefly, to 0.5 ml concentrated cells (as described for carbohy-

doi:10.1038/nbt.3865

drate analysis above), 800 µl of 6 N HCl containing 200 µl/ml thioglycolic acid, 10 µl of β-mecraptoethanol and 200 µl of 2 mM norvaline (internal standard) was added, and the vortexed sample was incubated at 110 °C for 24 h. Samples cooled to room temperature were centrifuged at 1,500*g* for 1 min, and a 50-µl aliquot was transferred to a fresh 2 ml GC vial. Aliquots were derivatized and analyzed by GC–MS according to the EZ:faast manual and ref. 44. This method allowed for the quantification of alanine, glycine, valine, leucine, isoleucine, proline, aspartate and asparagine, methionine, glutamate and glutamine, phenylalanine, lysine, tyrosine and cysteine. Tryptophan, threonine, serine, arginine and histidine were excluded. Volumetric concentrations of each detected hydrolyzed amino acid were converted to the carbon content present in that amount. These values were summed and normalized to TOC, as described for total carbohydrates above, thus yielding an estimate of carbon allocated to protein. However, when referring to FAME/TOC (as in **Fig. 1e**), values in grams FAME per gram TOC are presented.

Light-saturating photosynthesis capacity (P_{max}) and chlorophyll determination. P_{max} (based on ^{14}C incorporation) was measured in samples collected from cultures in semicontinuous growth mode 30 min before daily dilutions. Chlorophyll extractions were performed as previously described[45]. 0.5-ml aliquots of culture were placed into 2.0-ml screw-top tubes and centrifuged at 21,000*g* for 10 min. The supernatant was discarded, and 1.0 ml of 1:1 (vol/vol) dimethyl sulfoxide and acetone was added to each tube. Samples were agitated for 3 min in a prechilled block in a Mini-Beadbeater-96 (BioSpec Products) and then centrifuged for 10 min at 21,000*g*. Extracts were measured in disposable UV–vis cuvettes (BrandTech, cat. no. 759165) on a SpectraMax M2e (Molecular Devices) at 663 nm and 720 nm and quantified with the following equation: chlorophyll *a* (in µg ml^{-1}) = $(A_{663} - A_{720}) \times 20.15$ (ref. 45). Samples were normalized to 5 µg ml^{-1} chlorophyll *a* by centrifugation of cultures at 2,100*g* for 5 min, discarding of the supernatant and resuspension of the cells in the appropriate volume of standard medium (SM-NO_3^-) supplemented with 0.5 g l^{-1} sodium bicarbonate. Cultures were adapted to low light (~10 µmol photons m^{-2} s^{-1}) for approximately 1 h before P_{max} measurements, and 2-ml aliquots were placed into 7-ml glass scintillation vials (VWR, cat. no. 66022). A 20.4 µCi ml^{-1} working stock of ^{14}C-labeled sodium bicarbonate was prepared from a 1 mCi ml^{-1} stock (PerkinElmer, cat. no. NEC086H005MC) diluted in freshly prepared 3.57 mM sodium bicarbonate (Sigma-Aldrich, cat. no. S5761). 70 µl of the ^{14}C working stock was added to the 2.0 ml of diluted culture and mixed well with a multichannel pipette. 70 µl of each sample was added to a new set of 7-ml scintillation vials containing 100 µl of phenethylamine (Sigma-Aldrich, cat. no. 128945), filled with 4 ml of Ultima Gold AB (PerkinElmer, cat. no. 6013309) and capped immediately. An additional 1-ml aliquot of each sample was placed into another set of vials and acidified immediately with 2 N HCl. The remaining 1 ml of each samples was exposed to 2,500 µmol photons m^{-2} s^{-1} of white light for 10 min at 25 C. Samples were then immediately acidified with 2 N HCl and allowed to off-gas overnight. The following day, 4 ml of Ultima Gold AB was added to each acidified sample, and all samples were measured with an LS6500 scintillation counter (Beckman Coulter). Quantification was performed with equations from Littler and Arnold[46].

Western blotting. Cell pellets corresponding to 1.5×10^9 cells were resuspended in 500 µl buffer consisting of 125 mM Tris, pH 8.8, 10% glycerol and 2% SDS. Zirconium beads were added to the cell slurry, and samples were shaken twice for 3 min on a Mini Beadbeater (BioSpec Products). Lysates were boiled, incubated at 90 °C for 10 min and centrifuged, and the supernatant was diluted with 4× Laemmli sample buffer. Total protein extracted was determined with an RC DC protein assay kit (Bio-Rad). A total of 40 µg of protein was loaded per well onto a Tris-glycine (4–20%) denaturing gel. FAS proteins and ACCase were detected with antibodies raised by ProSci (peptide sequences in **Supplementary Table 7**). The primary antibody to NR was obtained from Agrisera (cat. no. AS08310; validation on manufacturer's website). iBind western blotting devices (Thermo Fisher Scientific) were used to incubate the blots with primary and secondary antibodies (anti-mouse AP, Novex, cat. no. A16081, 1:500 dilution). Immunosignals were detected with a colorimetric BCIP/NBT kit from Invitrogen (cat. no. 002209).

Transmission electron microscopy. Samples were pelleted and resuspended in fixing solution consisting of 2.5% glutaraldehyde, 2.5% paraformaldehyde and 0.1 M sodium cacodylate, pH 7.4. Transmission electron micrographs were captured by the Michigan State University Center For Advanced Microscopy with a high-resolution JEOL100 CXII microscope at an accelerating voltage of 100 kV.

Statistical analysis. No statistical methods were applied to predetermine sample size. Sample sizes were chosen on the basis of industry standards and experience. Differences between the performances of engineered strains and controls were analyzed with box-and-whisker plots showing median productivity, maximum and minimum data points, and first and third quartiles, as indicated in the corresponding figure caption. Differential expression analysis for RNA-seq experiments was performed with the R package edgeR[37]. For TF identification, FDR <0.01 was considered significant. Gene set enrichment testing was performed with goseq in R, and an FDR-adjusted *P* value <0.05 was considered significant.

Data availability. Sequence reads for all RNA-seq experiments described in this work have been deposited in NCBI's sequence read archive under accession number SRP096211. All other data supporting the findings of this study are available from the corresponding authors upon reasonable request.

35. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
36. Li, B. & Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
37. McCarthy, D.J., Chen, Y. & Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
38. Young, M.D., Wakefield, M.J., Smyth, G.K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).
39. Shen, B. *et al.* Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* **23**, 720–723 (2013).
40. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
41. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
42. Bligh, E.G. & Dyer, W.J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).
43. Ruiz-Matute, A.I., Hernández-Hernández, O., Rodríguez-Sánchez, S., Sanz, M.L. & Martínez-Castro, I. Derivatization of carbohydrates for GC and GC-MS analyses. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **879**, 1226–1240 (2011).
44. Kaspar, H. *Amino acid analysis in biological fluids by GC-MS.* Dr. rer. nat. thesis, Univ. Regensburg (2009).
45. Lichtenthaler, H.K. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. *Methods Enzymol.* **148**, 32 (1987).
46. Littler, M.M. & Arnold, K.E. In *Handbook of Phycological Methods. Ecological Field Methods: Macroalgae* (eds. Littler M.M. & Littler, D.S.) 349–375 (Cambridge University Press, 1985).